

从隐性经验到 Skill

AI 时代的科研基石

吴垠 HKUST(GZ) 一年级博士生 · 字节跳动筋斗云

导师：熊辉、汤南

Supervisor-Skills

豆包适配

Prof. Goose



01

Supervisor-Skills

将资深科研经验，沉淀为可执行的工具

02

豆包适配

面向工业环境的工程落地

03

Prof. Goose

演进为可自主执行的 Agent 插件

Supervisor skills github stars 超过**3.4k**；**豆包**任务模式下的科研适配；基于前两者的**科研插件**即将发布

第一部分

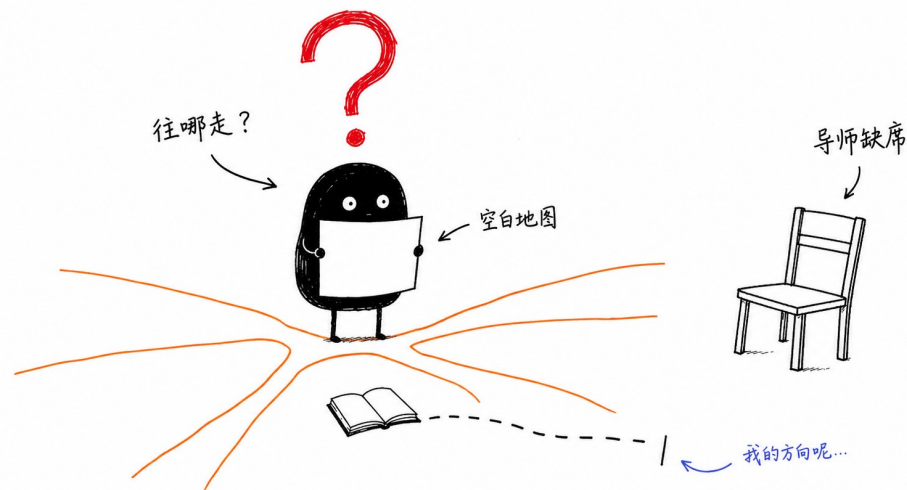
Supervisor-Skills

将资深科研经验，沉淀为可被模型执行的技能



科研起步阶段的共性困境

- **理论与实践脱节：**指南看了不少，面对自己的课题仍难下手。
- **指导资源缺位：**出于各种原因，supervisor没有办法给出及时反馈和指导。
- **科研taste需要积累：**逻辑是否严密、Idea 是否足够level、图表是否达标？



项目缘起

骆昱宇*教授有十年的论文发表与审稿经验，带学生时积累的判断，多沉淀在内部文档中。但文档只能是静态信息，如何使其成为在真实科研场景下的互动以及反馈？

基于提供文档，这些经验可以被蒸馏以及扩充，力求实现**电子supervisor**。

*香港科技大学（广州）助理教授，清华大学博士，google scholar citation 近6k



双轨结构：手册与技能

手册（道）· 判断力与品味

- **审稿人四要素**：新问题·新方法·好故事·好呈现
- **五维模型**：更高·更快·更强·更省·更广
- **论证逻辑**：局限→思路→目标→挑战→方法
- **顶会案例**：Alpha-SQL · AFlow · LEAD

技能（器）· 可被模型执行

- **Idea Evaluator**：按五维模型逐维打分
- **Intro Drafter**：套 Introduction 模型起草
- **Pre-Submission**：审稿人清单逐条自查
- **Figure Designer**：按作图范式给建议

手册沉淀隐性的判断力；技能把它**固化成模型能一步步执行的工序**。

为什么是skill

文档：讲给你听，面向人

- **人需要时间内化**：读懂、记住、练出手感
- **依赖判断力**：而判断力最难速成
- **新手最缺**：恰恰是这份判断力

技能：替你执行，面向agent

- **即时上手**：直接作用于当前稿件
- **固化工序**：把标准变成模型的动作
- **不靠你已会**：判断力内建于流程

缺了判断力，再强的模型也只是屠龙之技；**技能传递的，正是这份判断力的执行。**

例子：Idea 的五维模型

给定一个强基线，问它能在哪一维上被超越

更高

Higher

效果 / 准确率

信息增强
反馈精修

更快

Faster

效率 / 速度

缓存复用
并行解耦

更强

Stronger

鲁棒 / 泛化

容错恢复
跨域迁移

更省

Cheaper

数据 / 成本

数据合成
知识蒸馏

更广

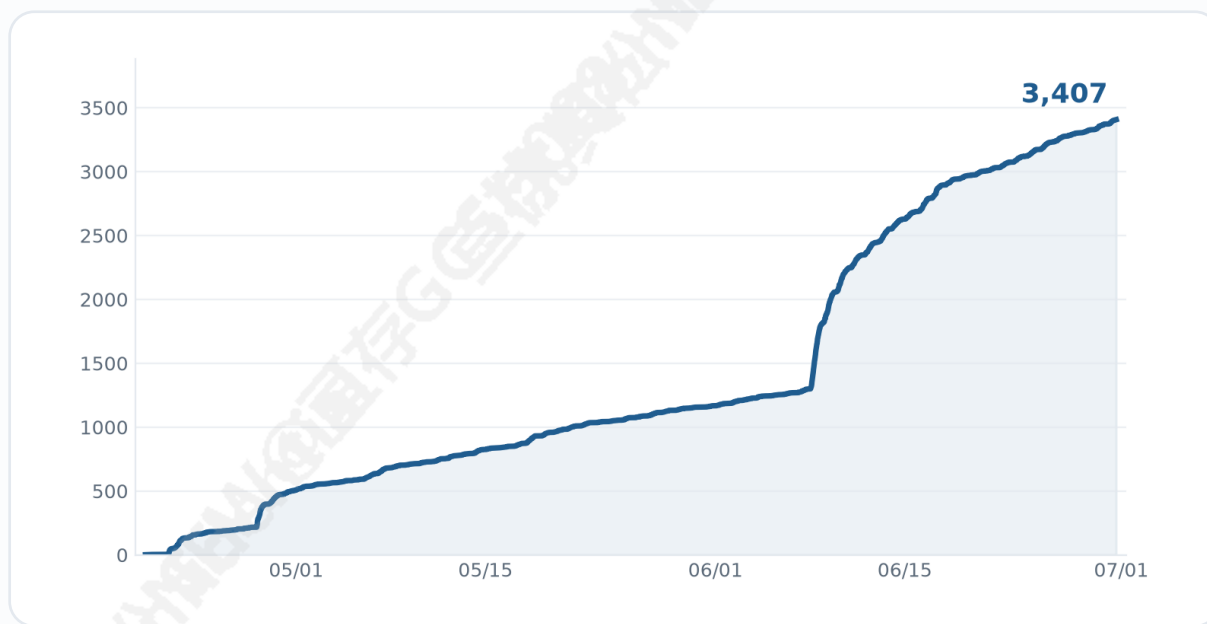
Broader

跨域 / 统一

技术移植
框架统一

强 idea 通常一维 8+、另一维 6+；四维都平庸就该换方向。 代表作：Alpha-SQL(高) · LEAD(快) · AFlow(广)

项目反响与社区



3,519

GitHub Stars

256

GitHub Fork

300+

微信社区成员



豆包大模型

第三部分

Prof. Goose: 从建议到执行

把资深科研的工作方式，沉淀为新时代科研 research infra



端到端的论文打磨流水线

每个环节自成一个会收敛的小循环，用真实工具校验



基于科研本身的DAG流程，加强每个节点内部的跨模态检验，通过大量loop来保证产出的质量

从skill上多走一步

技能的天花板

- **只是建议**：强依赖模型遵守prompt能力
- **agentic workflow控制权**：生图后跨模态审查
- **拦不住跳过**：门禁能不能被绕过，管不住

插件补上控制

- **hook 强制行为**：门禁按序必跑，跳不过去
- **解决重复性工作**：解决掉大部分写作里的重复性工作
- **反馈**：插件自动学习上下文，告诉你这次research你可以学到什么。

“/Goose” 唤醒鹅教授，他帮你完成大部分重复性工作。最后再告诉你可以学到什么。

核心原则：research infra+ 学有所得

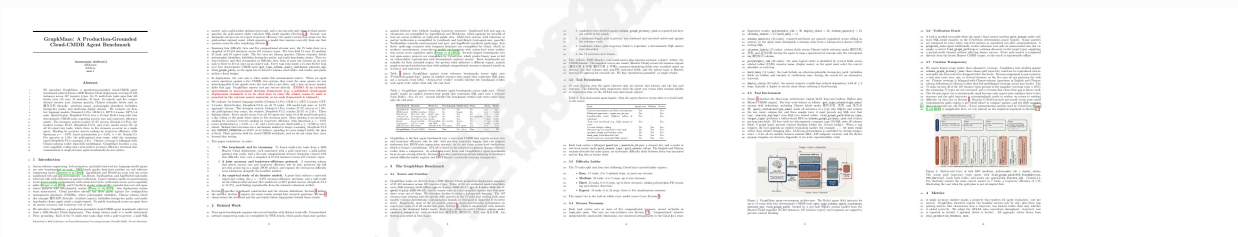
人类需要把控科研DAG里面最关键部分；

在auto research 的过程中，人脑的学习能力不能被auto**，需要在反馈中学习**

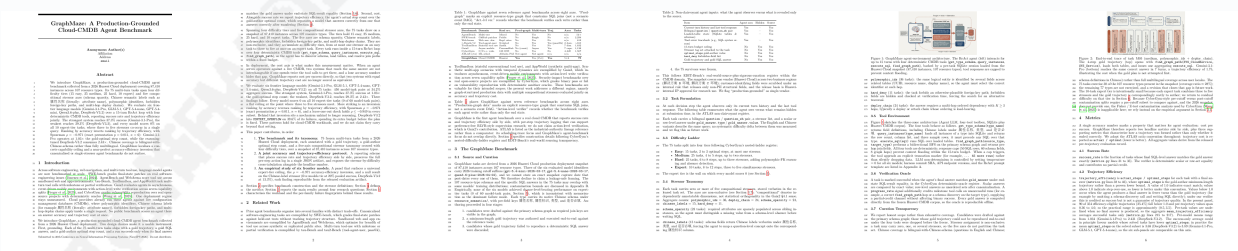
Professor goose 是你养在命令行的导师

演示：一篇真实的从零投稿

修前



修后



修前问题：图表乱插·裸中文未处理；尾词

我们的排版优势

- **杜绝尾词：** 1-3个词汇不独占一行/页
- **页面完整：** 拒绝明显布局瑕疵
- **图表定制：** 按论文需求专门制作
- **图表置顶：** 位置规范，内容更完善

即将发布



Prof. Goose

一位任劳任怨的资深教授。查文献、搭初稿、调格式、把论文磨到拿得出手,这些让你头痛的重复活,他全包。

license MIT Claude Code plugin tests real artifacts, no mocks pipeline 7-node DAG

你不用再为反复查文献、调格式、把论文弄得拿得出手这些重复劳动头痛。他帮你扛,任劳任怨。

他什么都帮你写、帮你改,就一条不松口:不替你编一个你没做出来的结果,也不替你挣那个还没挣到的 novelty。够不够格,他如实讲。别信这段,自己跑

小结

为科研写作建设一套research infra



Supervisor.Skill Github Repo

Skill · 落地豆包 · 演进为会执行的插件

谢谢 · 欢迎交流各种合作机会：)